International Journal of Computer Science Engineering
and Information Technology Research (IJCSEITR)
ISSN(P): 2249-6831; ISSN(E): 2249-7943
Vol. 4, Issue 4, Aug 2014, 9-14

TRANS STELLAR
• Journal Publications • Research Consultancy

AN EFFICIENT APPROACH FOR JOB SCHEDULING IN CLOUD COMPUTING

DEEPTI SARASWAT, DEEPTI SINGH & N. S. RAGHAVA

Information System, Delhi Technological University, Delhi, India

ABSTRACT

Cloud computing means computing over a network, where a program or applications may run on many connected computers at the same time. With the help of cloud computing, we can access several applications from different computers that are located at any other location. Job scheduling is one of the major issues in cloud computing. Many researchers are doing in-depth analyses of all possible scheduling algorithms which can give better results in terms of response time and overall efficiency of the system. In cloud computing environments,scheduling in virtual machinesis required for the execution of the job. Here we suggest a new scheduling algorithm that efficiently utilizes the resources of virtual machines and gives better results in cloud environments. This algorithm first properly analyses the resources of all virtual machines and then send job to that virtual machine which is most suitable for that job and also allows maximum number of jobs for execution.

KEYWORDS: Cloud Computing, Scheduling, Jobs, Algorithms

INTRODUCTION

Cloud Computing plays an important role for business institutions as well as research institutions, in the last few years. Cloud computingis a combination of various technologies like virtualization, distributed computing, networking, software and web services. Major components of cloud includes clients, datacenter and distributed servers. [1-2] It includes fault tolerance, high availability, scalability, flexibility, reduced overhead for users, reduced cost of ownership, ondemand services, etc. Cloud computing is a promising technology to provide on demand services according to the client's requirements within a promised time. Further, the cloud computing environment provides the facility to the users for accessing the shared environment of distributed resources. Cloud is a pay- go model where the consumers pay for the resources that they have utilized, which require to have highly available resources to service the requests on demand. Therefore, the complexity of managing the resources from the business perspective of the cloud service provider increases.

Scheduling is one of the major work performed in all the computing environments. To increase the efficiency of cloud computing systems, job scheduling is one the important task which requires massive attention from developers.This can lead to maximum profit for both the users and the cloud service providers. The main aim of scheduling algorithms in cloud computing systems is to properly utilize the processing units and reduce the execution time of job.Scheduling, one of the most famous optimization problems, plays a key role to improve flexibility and reliable systems. The main purpose is to schedule jobs to the adaptable resources in accordance with adaptable time, which involves finding out a proper sequence in which jobs can be executed under transaction logic constraints.

There are main two categories of scheduling algorithm.

- **Static Scheduling Algorithm:** This algorithm uses the concept of pipelining. Itfirst prefaces the required data and apply pipelining on different stages of task execution. This scheduling imposes less runtime overhead.

- **Dynamic Scheduling Algorithm:** In this algorithm, information about the job components/task is not known beforehand.

Thus the execution time of the task may not be known and the allocation of tasks is done on the fly as the application executes. Both have their own advantages and limitations. Dynamic scheduling algorithm has higher performance than static algorithm but has a lot of overhead compare to it.

SCHEDULING

In a distributed system, varieties types of scheduling algorithm has been used. We can also use those algorithms in environment of cloud by making certain changes and also verify them in our environment. The main motive of the job scheduling algorithm is to gain a high performance computing and the best system throughput. For providing job scheduling algorithm in cloud we cannot use them directly. Job scheduling algorithm can simply classify into two groups; Batch mode heuristic scheduling algorithms (BMHA) and online mode heuristic algorithms. In BMHA, first we make a queue of jobs and then grouped into a batch when they arrive into the system. So after a certain period of time, scheduling algorithm starts. Following are the examples of BMHA based algorithms; First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Min–Min algorithm and Max–Min algorithm [3-4]. In On-line mode heuristic scheduling algorithm, scheduling of jobs is done when they arrive the system. So, on-line mode heuristic scheduling algorithms are better to apply in a cloud environment. Most fit task scheduling algorithms (MFTF) are one of the examples of On-line mode heuristic scheduling

Algorithm

- **First Come First Serve Algorithm**

In this algorithm first they make a queue of jobs. Then they form a batch of jobs from where they select jobs for execution on the basis that which jobs comes first. It is simple and fast. But the disadvantage of this algorithm is that if a job have higher priority than it should wait for his turn which may be sometime not a feasible solution.

- **Round Robin Algorithm**

In the round robin scheduling, processes are dispatched in a FIFO manner, but a certain amount of CPU time is given to all process when their turn comes, that time is called a time-slice or a quantum. If a process does notcomplete before its time-slice expires, then the CPU is primped and given to the next process waiting in a queue. The preempted process is then placed at the end of the ready list.

- **Min–Min Algorithm**

In this algorithm,those processes are chosenfirst which requires a small time for execution.It means it gives higher priority to small jobs. But in this algorithm starvation for longer jobs may be possible.

- **Max – Min Algorithm**

In this algorithm, those processes are chosen first which requires a longer time for execution. It means it gives higher priorityto longer jobs. But in this algorithm starvation for smaller jobs may be possible.

- **Most Fit Task Scheduling Algorithm**

  In this algorithm task which fit best in the queue are executed first. This algorithm has a high failure ratio.

- **Priority Scheduling Algorithm**

In this algorithm, a priority is assigned to each job either internally or externally. On the basis of that priority jobs are selected for execution. Shortest job first (SJF) is one of the examples of priority scheduling algorithm. In this highest priority is given to those jobs that requires less time of CPU for execution. And if more than one job requires same CPUtime, then FCFS algorithm is applied [5-6]

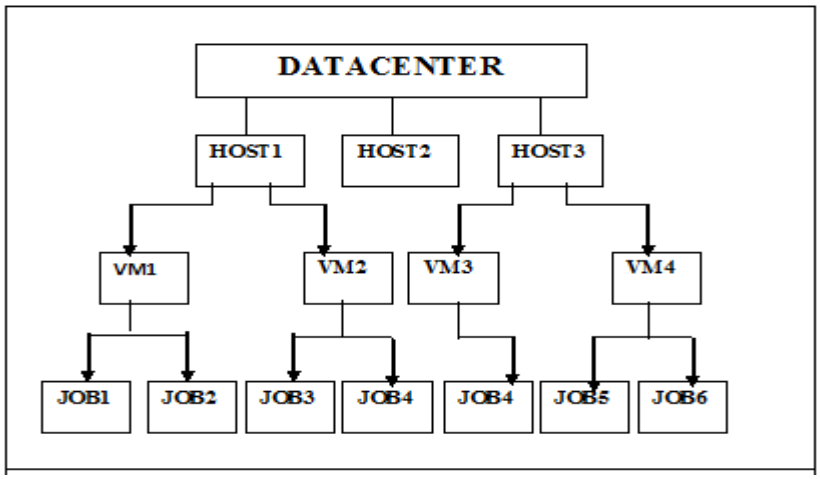

Figure 1: Process of Scheduling

Scheduling Process

The process of scheduling in the cloud is divided into three steps:-

- **The Resource is Discovering and Filtering:** Datacenter Broker discovers the resources present in the network system and collects information about their status.
- **Resource:** On the basis of certain parameters resources and tasks are selected.. This is deciding stage.
- **Task Submission:** Task is submitted to resource selected.

Proposed Algorithm

Step1

In this algorithm we consider that each virtual machine has its resources. We take here only four resources including hardware and software (CPU, main memory, hard disk, softwares) but in a real cloud environment, we can take any number of resources. Here virtual machine along with its resources are represented in the form of a matrix. Here is the example that describes the algorithm.

Table 1

	No of CPU	Main Memory	Hard Disk	Softwares
Vm1	3	5	4	9
Vm2	2	4	4	8
Vm3	2	4	2	1
Vm4	3	3	3	1

Step 2

Table 2

Job1	2	4	4	8

The above row represent the number of resources required by Job1.Here we have two options that either we allocate it to the first virtual machine or to the second virtual machine. Here we consider these two as a separate case:

Case 1

Suppose we allocate this cloudlet to first virtual machine.

Then we update the matrix by subtracting the total no of a particular resource available in virtual machines to the total no of resources required by that job. Here it is like that for spa(3-2) >0 if this term is greater than zero, then do this thing for all the resources otherwise neglect this virtual machine and move to the next virtual machine. Similarly for main memory (5-4) >0,hard disk (4-4) >0, softwares (9-8) >0.

Now our updated matrix becomes:-

Table 3

	No of CPU	Main Memory	Hard Disk	Softwares
Vm1	1	1	0	1
Vm2	2	4	4	8
Vm3	2	4	2	1
Vm4	3	3	3	1

Now suppose job2 comes

Table 4

Job2	3	5	4	9

So Here we have no virtual machine that has enough number of resources that is required by the job1. Hence In this case we are only able to execute only one job. Because we have no virtual machine which satisfy the resource requirement of job2.

Case 2

Suppose we allocate job1 to second virtual machine, then our resulting matrix becomes by following the same procedure that we have done for first cloudlet in the first case.

Table 5

	No of CPU	Main Memory	Hard Disk	Softwares
Vm1	3	5	4	9
Vm2	0	0	0	0
Vm3	2	4	2	1
Vm4	3	3	3	1

No request for job2

Table 6

Job2	3	5	4	9

Now in this case we are able to execute job2 also. I. e. We execute both the cloudlet efficiently. So from above two cases we conclude that we have to select that virtual machine which gives the minimum sum after the summation of all the differences that we have calculated for corresponding resources. If summation is same for more than one virtual machine, then allocated to any of the virtual machines.

Step 3

When the job is completely executed, then update the matrix of availability of resources. The above cases arise in any condition so we apply this algorithm for request of each job. This scheduling algorithm allows maximum number of jobs for execution.

CONCLUSIONS

Scheduling in cloud computing plays a very important role in establishing the cloud computing environment. The main aim of our scheduling algorithm is to manage the maximum number of jobs in a very efficient manner. With the help of this algorithm we can increase the resource utilization by executing number of jobs and also efficiency of this algorithm does not vary if the number of resources increases. Also the time required forthe computation is very less it just require the basic calculation.

REFERENCES

1. Pinal Salot M.E, Computer Engineering, Alpha College of Engineering, Gujarat, India (2013) A survey of various scheduling algorithmsin a cloud computing environment.
2. Poonam Devi Department of CSE BITS College, BHIWANI, MDU, India (2013)BITS College. Implementation of Cloud Computing By Using Short Job Scheduling.
3. Sujit Tilak,2Department of COMPUTER, PIIT, New Panvel, Maharashtra, India (2012).A Survey of Various Scheduling Algorithms in Cloud Environment
4. Yogita Chawla 1,2Pune University, G.H Raisoni College of Engg & Mgmt,Gate No.: 1200 Wagholi, Pune – 412207 (2012) A Study on Scheduling Methods in CloudComputing.
5. Jairam Naik K. Research Scholar, Faculty of Computer Science & Engineering, JNT University, Hyderabad, India (2012) Scheduling Tasks on Most Suitable Fault tolerant Resource for Execution in Computational Grid.
6. Satchel Patel (2013) Priority Based Job Scheduling Techniques In Cloud Computing: A Systematic Review